5-2008

Dow Jones Index, GARCH(1,1) and Change-Points

Tharanga Wickramarachchi
Clemson University, twickra@clemson.edu

DOW JONES INDEX, GARCH(1,1) AND CHANGE-POINTS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mathematical Sciences

by
Tharanga Wickramarachchi
May 2008

Accepted by:
Dr. C.M. Gallagher, Committee Chair
Dr. A. Aue, Co-Advisor
Dr. R.B. Lund
Dr. K.B. Kulasekera

Abstract

Many econometric time series data sets, such as log returns of stocks, exhibit evidence of the so called stylized facts. Namely it is generally observed that the data itself is uncorrelated with heavy tails, but the squared data has significant autocorrelation. For such data sets, there appears to be little or no linear information in the past about the future values of the series. Thus the class of Autoregressive Integrated moving average models (ARIMA) are not appropriate. However, there does in general appear to be information in past values of the squared data about future values of the squared data. This allows for modeling of the conditional variance as a function of the observed past. One choice for a class of models able to incorporate the stylized facts are the generalized autoregressive conditional heteroskedastic (GARCH) family. In practice the GARCH(1,1) process is used most often.

This work reviews what is known about the GARCH(1,1) model and investigates the appropriateness of a GARCH(1,1) model for daily Dow Jones Index stock returns (DWJ). A GARCH(1,1) model is fit to the DWJ series. Based on visual inspection of the return data, there may be one or more changepoints in the process governing the data. We use a recent test proposed by Berkes, Horváth, and Kokoszka in order to locate possible change-points in the DWJ data. Although the test is designed to detect a single changepoint in a GARCH process, the test is applied sequentially in an attempt to find mutliple changepoints. It is found that a single GARCH(1,1) model cannot be fitted for the DWJ series from January 1997 through December 2006. Sequential changepoint testing indicates a single changepoint in the data. In order to assess the reliability of these conclusions, a simulation study is used to investigate the properties of the changepoint test. The change-point test performs quite well in terms of type I error, but its power is small. Overall a better test for change-point detection in GARCH(1,1) processes would be welcome.

Dedication

I dedicate this work to my loving family. It is their love, support and encouragement that made this work complete.

Acknowledgments

I would like to express my heartiest gratitude to my advisor Dr. C.M. Gallagher for his guidance and support. His thoughtfulness and new ideas inspired me and helped me immensely in making this a success. My sincere thanks to Dr. A. Aue, for sharing his world of experience and mathematical insights during the course of this work and also many thanks to Dr. R.B. Lund for his insightful suggestions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the early stages of time series analysis and econometric modeling, constant variance assumptions played a crucial role. Most statistical tools were geared towards linear processes such as the highly popular autoregressive moving average (ARMA) sequences. Since the 1980s, however, a vast theoretical and applied framework in nonlinear econometric modeling has emerged, which was triggered by the introduction of the autoregressive conditionally heteroskedastic (ARCH) process in the seminal work of Engle (1982). Engle, who recently was awarded the Nobel Price in economics, has suggested that financial data may be better described by assuming that the conditional variance of the underlying process changes as a function of past errors.

Some of the pioneering steps were taken by Engle (1982) (6), Engle (1983) (7), Engle and Kraft (1983) (8), Weiss (1984) (14), Coulson and Robins (1985) (4), Engle, Lilien and Robins (1985) (9), and Domowitz and Hakkio (1985) (5). These papers put a focus on the "introduction of a rather arbitrary linear declining lag structure in the conditional variance equation", which was identified as a common feature in a variety of financial applications (see Bollerslev, 1986 (3)). The need for this flexibility has eventually lead to the introduction of the generalized autoregressive conditionally heteroskedastic (GARCH) model by Bollerslev (1986), which is an extension of the original ARCH process of Engle (1982).

In this work we will focus on the popular GARCH(1,1) model. In the remainder of this chapter we review what are known as "stylized facts" in the financial econometrics literature and introduce the GARCH(1,1) model in light of these facts. We review all the theoretical considerations of the GARCH(1,1) model in chapter 2. In chapter 3 we describe a test proposed by Berkes, Horváth, and

Kokoszka (2004) which can be used to detect structural change in a GARCH sequence. In chapter 4 we apply the theory to fit GARCH(1,1) models to the daily Dow Jones Index stock returns (DWJ). The changepoint test indicates strong evidence of at least one changepoint in the GARCH structure of the Dow Jones Index returns. Thus one GARCH(1,1) model cannot be used to describe the DWJ series. However, when applied sequentially the changepoint test indicates no further division of the data is necessary. We investigate the reliability of these conclusions via a simulation study which investigates the type I error and power propeties of the changepoint test. It appears that the power of the test can be low. Thus there may in fact be more than one changepoint in the GARCH structure of the DWJ series.

Suppose that we have an arbitrary discrete-time financial data process (P_t) which we observe at regularly spaced times $t = 0, 1, \ldots$. It has been pointed out (e.g., see Mikosch, 2000) (12) that a great variety of econometric time series simultaneously display a set of common properties after the log-transformation

$$x_t = \log P_t - \log P_{t-1} = \log\left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right), \quad \text{where} \quad t = 0, 1, 2, ... \tag{0.1}$$

is applied. These properties, however, depend on the time scale at which observations are being made. If this scale is too small, for example, then then changes in (P_t) will occur rather infrequently, in turn often causing x_t to be zero. Commonly, the process (P_t) is referred to as the price process, while (x_t) is called the log returns of (P_t) for the following reason. Assuming that the changes $P_t - P_{t-1}$ are small compared to the values of P_t, a Taylor expansion for $\log(1 + x)$ yields that $x_t \approx (P_t - P_{t-1})/P_t$.

1.1 Stylized Facts

Samples of log-returns exhibit the following stylized facts.

1. Distribution and tails.

- The sample mean of the log-returns is close to zero; their sample variance is of the order 10^{-4} and smaller. Since the price changes are generally small, it is an obvious property.
- A density plot of the log-returns shows that their distribution is roughly symmetric in its center, sharply peaked around zero, with heavy tails on both sides.

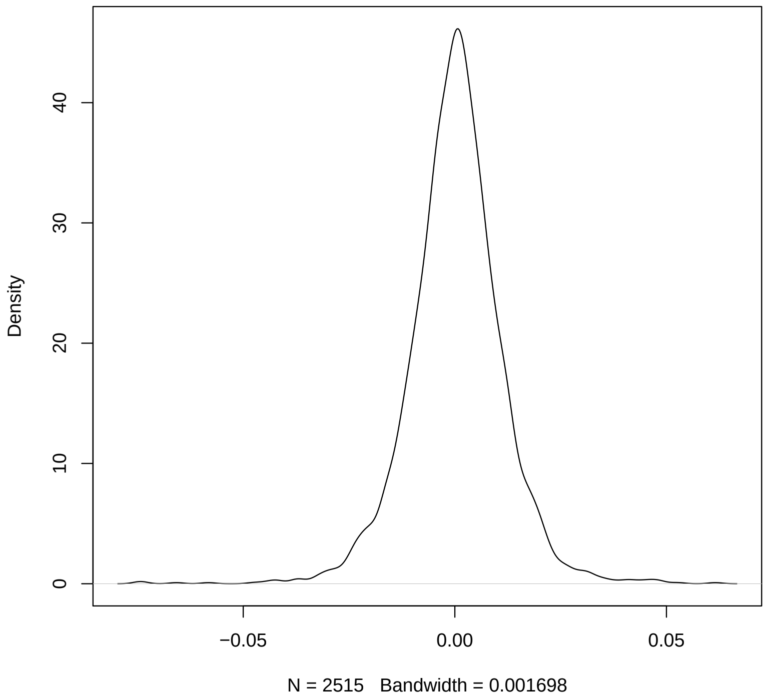

Figure 1.1: *Density Plot for Log-Returns of Dow Jones Index Series from Jan-1997 through Dec-2006*

The density plot in Figure (1.1) for the log-returns of the Dow Jones stock index (DWJ) confirms the stylized fact mentioned at last. We can clearly see a symmetric distribution and it is sharply peaked around zero with relatively heavy tails.

2. Dependence, autocorrelations and clusters of extremes.

- The sample autocorrelation function (sample ACF) $\rho_{h,x}$ is negligible at all lags. (A possible exception is the first lag. However the estimated value is usually small in absolute value as well.)

The sample ACF in Figure (1.2) for the log-returns of DWJ series shows that almost all auto-correlation values are negligible, which supports the claim made above.

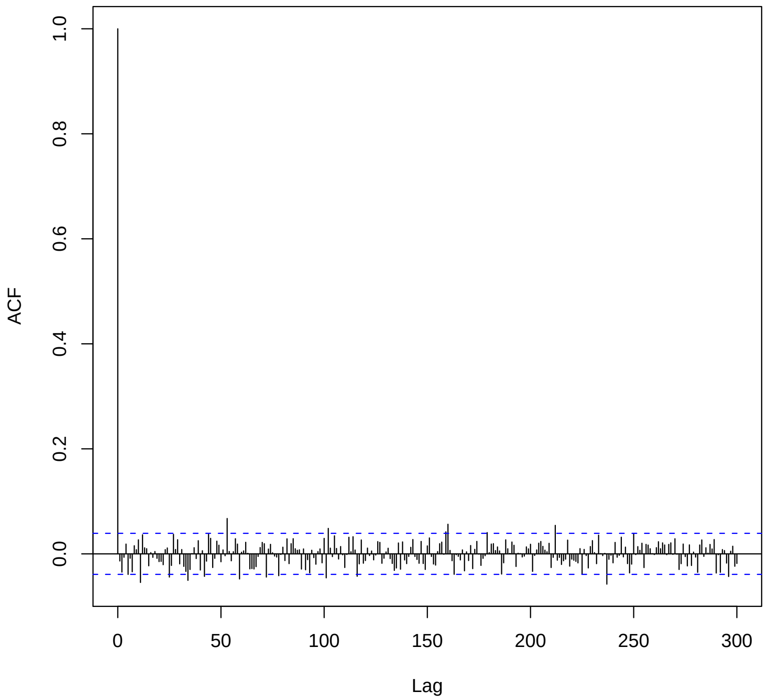

Figure 1.2: *Sample ACF for Log-Returns of Dow Jones Index Series from Jan-1997 through Dec-2006*

- The sample ACFs $\rho_{n,|x|}$ of the absolute values $|x_t|$ (referred to as the absolute log-returns) and ρ_{n,x^2} of the squares, x_t^2, are different from zero for a large number of lags and stay almost constant and positive for large lags.

The plots in Figure (1.3) confirm the fact that the sample ACFs for both squares and absolute values are different from zero for a large number of lags. It is also noticeable that ACF values remain positive for a large number of lags.

- The large and small values in the log-return sample occur in clusters.

3. Aggregational Gaussianity.

- The distribution of log-returns over longer periods of time (such as a month, half a year, a year) is closer to the normal distribution than for hourly or daily log-returns.

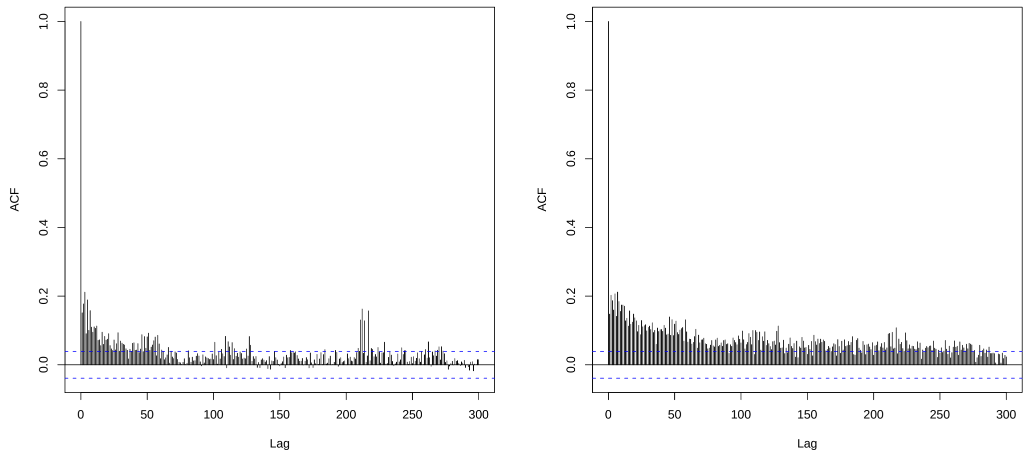

Figure 1.3: *Sample ACFs for the Squared log-returns (left) and absolute log-returns (right) of the Dow Jones Index*

Here, we will focus on the widely popular GARCH(1,1) model and how it captures all or some of these stylized facts. In the following section, we give its definition and outline the contents of all chapters to come.

1.2 The definition of a simple GARCH process

Denote the set of integers by $\mathbb{Z}$.

Definition 1.2.1 *A real-valued discrete-time stochastic process* $\{x_t : t \in \mathbb{Z}\}$ *is called a generalized autoregressive conditionally heteroskedastic process of order* $(1,1)$*, shortly GARCH(1,1), if it satisfies the equations*

$$x_t = \sigma_t \varepsilon_t, \tag{2.2}$$

$$\sigma_t^2 = \omega + \alpha \sigma_{t-1}^2 + \beta x_{t-1}^2, \tag{2.3}$$

where $\{\varepsilon_t : t \in \mathbb{Z}\}$ *is a sequence of independent, identically distributed (henceforth iid) random variables with* $E[\epsilon_t] = 0$ *and* $E[\epsilon_t^2] = 1$*, and* $\omega > 0$*,* $\alpha \geqslant 0$ *and* $\beta \geqslant 0$ *are real-valued coefficients.*

It is easy to show by starting from equation (2.2) and subtracting σ_t^2 from both sides of the equation, that the variance structure is equivalent to an ARMA(1,1) process. (This is shown in detail in the proof of Theorem 2.2.3 in Chapter 2.) In linear time series analysis, autoregressive moving average (ARMA) processes are widely used (see, for example, Brockwell and Davis, 1991, for a detailed exposition).

Definition 1.2.2 *A real-valued discrete-time stochastic process* $\{x_t : t \in \mathbb{Z}\}$ *is called autoregressive moving average of order* $(1,1)$*, shortly ARMA(1,1), if it satisfies the equations*

$$x_t - \phi x_{t-1} = \varepsilon_t + \theta \varepsilon_{t-1}, \tag{2.4}$$

where $\{\varepsilon_t : t \in \mathbb{Z}\}$ *is a sequence of iid random variables with* $E[\varepsilon_t] = 0$ *and* $E[\varepsilon_t^2] = \sigma^2$*. The (linear) polynomials* $1 - \phi z$ *and* $1 + \theta z$ *are referred to as the autoregressive and moving average polynomial, respectively.*

A set of data is simulated from a GARCH(1,1) model with parameters $\omega = 1 \times 10^{-6}, \alpha = 0.8$ and $\beta = 0.1$ in order to study how well the GARCH(1,1) model captures some of the stylized facts specified in Section 1.1. The density plot in Figure (1.4) corresponding to the simulated series also shows a symmetric distribution peaked around zero with heavy tails. Therefore it captures the stylized fact, which describes the density plot.

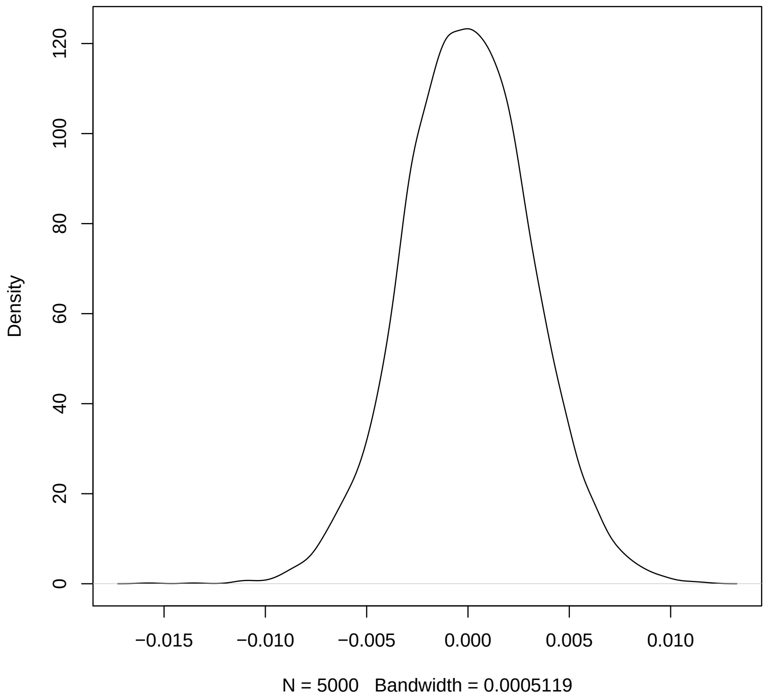

Figure 1.4: *Density plot for the GARCH(1,1) series with* $\omega = 1 \times 10^{-6}$*,* $\alpha = 0.8$*,* $\beta = 0.1$

The sample ACF for the simulated series in Figure (1.5) verifies the fact that the auto-correlation values are negligible at all lags. That provides the evidence that the GARCH(1,1) model follows the stylized fact corresponding to the ACF of the series.

Now we study the ACFs for the squares and absolute values of the simulated series to see how well it captures the corresponding stylized fact.

It can be seen that the auto-correlation values in Figure (1.6)are different from zero and remain positive for a large number of lags for both squares and absolute value series. Therefore it is clear that the GARCH(1,1) model captures most of the stylized facts specified in Section 1.1. There are a number of questions that immediately arise from Definition 1.2.1. In particular, we are going to address the following ones here.

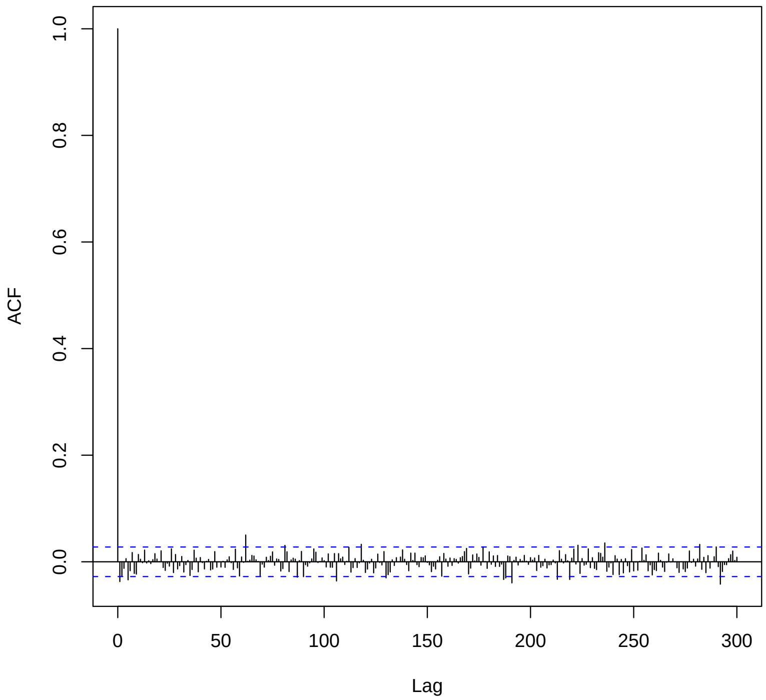

Figure 1.5: *Sample ACF for the GARCH(1,1) series with* $\omega = 1 \times 10^{-6}$, $\alpha = 0.8$, $\beta = 0.1$

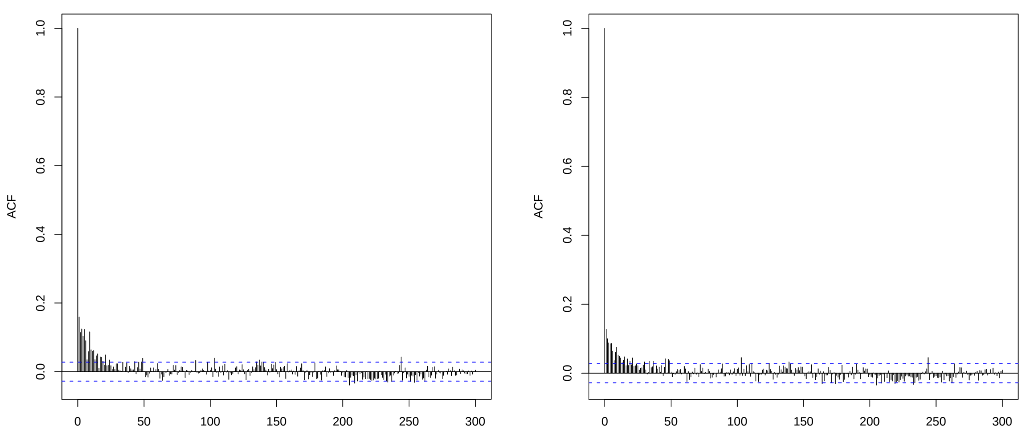

Figure 1.6: *Sample ACFs for the squares (left) and absolute (right) values of the GARCH(1,1) series with* $\omega = 1 \times 10^{-6}$, $\alpha = 0.8$, $\beta = 0.1$

- How can we characterize sequences $(x_t : t \in \mathbb{Z})$ that solve the system of equations specified in (2.3)? How can we compute moments and other theoretical features from the definition? Answers to these questions shall be provided in Chapter 2, which contains a collection of probabilistic results on the structure of GARCH(1,1) processes.

- How does one estimate efficiently the unknown parameters ω, α and β in (2.3)? The statistical analysis of GARCH(1,1) processes will be part of Chapter 3.

- How can the structural changes be identified? One way of accomplishing this task will be discussed in Chapter 3.

- In Chapter 4 we will discuss how well GARCH(1,1) models can be fitted to real financial data such as the log-returns of the Dow Jones Stock Index.

Chapter 2

The Probabilistic Properties of GARCH(1,1) Processes

In this chapter, we discuss the basic but most important probabilistic properties of the GARCH(1,1) process given in Definition 1.2.1. We introduce the concepts of strict and weak (or, second-order) stationarity in Section 2.1, and provide necessary and sufficient conditions for the existence of strictly and weakly stationary solutions to the system of equations in (2.2) and (2.3). In Section 2.2, we derive conditions for the finiteness of moments of a stationary GARCH(1,1) process.

2.1 Stochastic difference equations and their solutions

In a wider context GARCH models fall into the category of random sequences which are given by *stochastic difference equations* (see, among many others, Furstenberg and Kesten, 1960; and Mikosch and Stărică, 2000). In this section, we will demonstrate how the GARCH(1,1) equations (2.2) and (2.3) can be solved in order to obtain the necessary and sufficient conditions that guarantee the existence of a unique strictly (weakly) stationary solution. To this end, we start with the following definition.

Definition 2.1.1 *(a) Strictly stationary sequence*

A sequence $\{x_t : t \in \mathbb{Z}\}$ is strictly stationary if $(x_1, \ldots, x_n)$ and $(x_{1+h}, \ldots, x_{n+h})$ have the same joint distributions for all integers h and $n > 0$.

(b) *Weakly stationary sequence*

A sequence $\{x_t\}$ *is weakly stationary if,*

(i) $E[x_t] = \mu_x(t)$ *is independent of* t,

and

(ii) $\gamma_x(t+h,t) = \mathrm{Cov}(x_{t+h}, x_t)$ *is independent of* t *for each* h.

Let $\{x_t : t \in \mathbb{Z}\}$ be a GARCH(1,1) process according to Definition 1.2.1. As a starting point in deriving stationary solutions, we repeatedly use equations (2.2) and (2.3) to obtain,

$$\begin{aligned}
\sigma_t^2 &= \omega + \alpha\sigma_{t-1}^2 + \beta x_{t-1}^2 \\
&= \omega + (\alpha + \beta\epsilon_{t-1}^2)\sigma_{t-1}^2 \\
&= \omega + (\alpha + \beta\varepsilon_{t-1}^2)(\omega + \alpha\sigma_{t-1}^2 + \beta x_{t-2}^2) \\
&= \omega + (\alpha + \beta\varepsilon_{t-1}^2)\left(\omega + [\alpha + \beta\varepsilon_{t-2}^2]\sigma_{t-2}^2\right) \\
&= \omega \sum_{i=1}^{n} \prod_{j=1}^{i-1} M_t(j) + \sigma_{t-n}^2 \prod_{j=1}^{n} M_t(j), \qquad (1.1)
\end{aligned}$$

where an empty product is set to equal one and where

$$M_t(j) = \alpha + \beta\varepsilon_{t-j}^2, \qquad j \geq 1.$$

It is now apparent from equation (1.1), that the necessary and sufficient conditions for the existence of stationary solutions will follow from the existence of the random variable

$$S^2 = \sum_{i=1}^{\infty} \omega \prod_{j=1}^{i-1} (\alpha + \beta\varepsilon_{-j}^2). \qquad (1.2)$$

This can be heuristically explained by letting $n \to \infty$ in an appropriate sense (to be discussed below) in equation (1.1): The necessary and sufficient conditions must ensure that the sum converges and that the product does not contribute to the limit. Note also that, in the case that the limit exists, we have $S^2 = \sigma_0^2$. Our first result concerns strict stationarity. In the following, we assume that

$$E[\log^+ |\alpha + \beta\varepsilon_0^2|] < \infty. \qquad (1.3)$$

Moreover, we call a strictly stationary solution $\{x_t : t \in \mathbb{Z}\}$ of (2.2) and (2.3) *non-anticipative* if x_t is independent of $\{\varepsilon_s : s > t\}$ for all $t \in \mathbb{Z}$.

Theorem 2.1.1 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2) and (2.3) that satisfies (1.3). Then, $\{x_t : t \in \mathbb{Z}\}$ admits a unique non-anticipative strictly stationary solution if and only if*

$$E[\log|\alpha + \beta\varepsilon_0^2|] < 0. \tag{1.4}$$

It is clear that Theorem 2.1.1 is proved if we can show that the random sum S^2 in (1.2) is convergent almost surely if and only if equation (1.4) holds. Since it can be shown that the product of the equation (1.2) disappears as $n \to \infty$, we only need to show that S^2 converges almost surely. The proof is given in Section 2.3 below. Here, we continue with the characterization of weakly stationary solutions. Note that (1.4) is satisfied whenever $\alpha + \beta < 1$ (As a result of Jenkins Inequality). In a similar fashion, we can now state the counterpart of Theorem 2.1.1 for the weakly stationary case.

Theorem 2.1.2 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2) and (2.3). Then, $\{x_t : t \in \mathbb{Z}\}$ admits a unique weakly stationary solution if and only if $\alpha + \beta < 1$.*

For the proof of Theorem 2.1.2, we will need to investigate more closely moment properties of the GARCH(1,1) process. This will be done in the next section, while the proof itself is deferred to Section 2.3. It was proved that satisfying the condition $\alpha + \beta < 1$ is sufficient for strict stationarity, but the converse does not always need true.

2.2 Moments and dependence structure

In this subsection, we present necessary and sufficient conditions for the finiteness of moments of a strictly stationary GARCH(1,1) sequence. Furthermore, we derive the first and second moments as well as the auto-covariance function of the GARCH(1,1) sequence and its squares.

Theorem 2.2.1 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2) and (2.3) which satisfies equations (1.3) and (1.4). Let $\nu > 0$. If*

$$E\left[|\alpha + \beta\varepsilon_0^2|^\nu\right] < 1, \tag{2.5}$$

then $E[|S^2|^\nu] < \infty$. Conversely, if $E[|S^2|^\nu] < \infty$ for some $\nu > 0$, then (2.5) holds.

When $E\left|S^2\right|^{\nu} < \infty$, we immediately get $E\left|\sigma_t^2\right|^{\nu} < \infty$ (the product term does not contribute to the limit of equation 1.1). Since σ_t^2 and ε_t^2 are independent and $E\left[\varepsilon_t^2\right] < \infty$, we can derive that if $E\left[\left|S^2\right|^{\nu}\right] < \infty$, then $E\left[|x_t|^{\nu}\right] < \infty$. In the next step, we derive the first and second moments and the auto-covariance function of a GARCH(1,1) process.

Theorem 2.2.2 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2) and (2.3) which satisfies equations (1.3) and (1.4). If $E[|\varepsilon_t|^{\nu}] < \infty$ and if (2.5) is satisfied for some $\nu \geq 1$, then it holds that*

$$E[x_t] = 0 \qquad \text{and} \qquad E[x_t^2] = \frac{\omega}{1 - \alpha - \beta} \qquad \text{for all } t \in \mathbb{Z}. \tag{2.6}$$

Furthermore,

$$\operatorname{Cov}(x_t, x_s) = 0 \qquad \text{for all } t, s \in \mathbb{Z},\ t \neq s. \tag{2.7}$$

Theorem 2.2.3 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2) and (2.3) which satisfies equations (1.3) and (1.4). If $E[|\varepsilon_t|^{\nu}] < \infty$ and if (2.5) is satisfied for some $\nu \geq 2$, then it holds that*

$$\operatorname{Cov}(x_t^2, x_{t+h}^2) = \gamma_{x^2}(h) = \begin{cases} \sigma_v^2\left[1 + \frac{\beta^2}{1-(\alpha+\beta)^2}\right] & h = 0 \\ \sigma_v^2\left[1 + \frac{\beta(\alpha+\beta)}{1-(\alpha+\beta)^2}\right]\beta & h = 1 \\ (\alpha+\beta)^{h-1}\gamma_{x^2}(1) & h > 1 \end{cases} \tag{2.8}$$

Theorem (2.2.2) and Theorem (2.2.3) tell us, even though the x_ts are uncorrelated, squares (x_t^2) are correlated and it decays exponentially. This feature is also stated in Section 1.1. An illustration of these facts for DWJ returns data is presented in Figure 2.1.

Figure 2.1 shows the sample ACFs for a simulated series (left) and its squares (right), which are obtained from a GARCH(1,1) process with parameters $\omega = 1 \times 10^{-6}$, $\alpha = 0.8$, $\beta = 0.1$. It clearly shows that the auto-correlation for log-returns are negligible at most of the lags and for squared log-returns it drops down to zero exponentially.

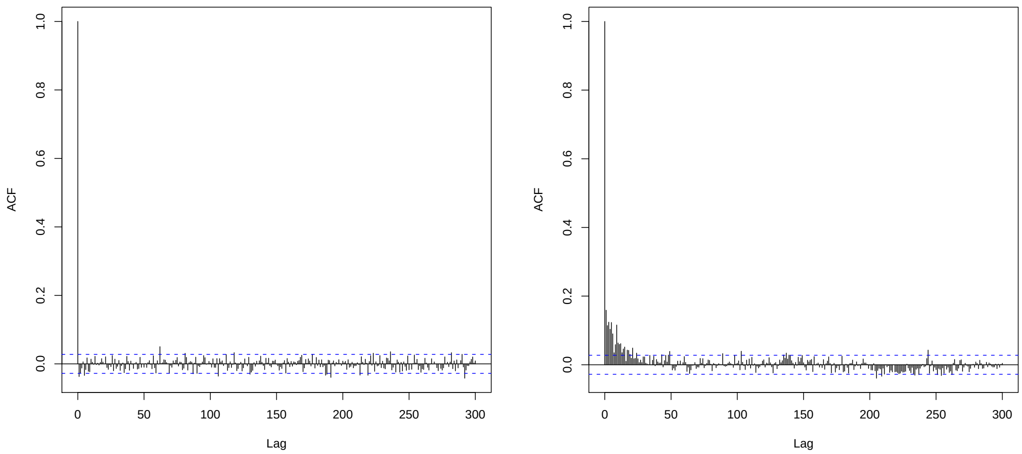

Figure 2.1: *Sample ACFs for the series (left) and squares (right) Values of the GARCH(1,1) Process with* $\omega = 1 \times 10^{-6}$, $\alpha = 0.8$, $\beta = 0.1$

2.3 Mathematical proofs

Most of the proofs are rewritten specifically for a GARCH(1,1) model considering more general proofs given in different papers.

Proof of Theorem 2.1.1

In order to prove theorem 2.1.1, the equation 1.1 is rewritten as follows.

$$\begin{aligned}\sigma_t^2 &= \omega \sum_{i=1}^{n-1} \prod_{j=1}^{i-1} M_t(j) + \sigma_{t-n}^2 \prod_{j=1}^{n} M_t(j) \\ &= S_n^2 + R_n^2\end{aligned}$$

Where

$$S_n^2 = \omega \sum_{i=1}^{n-1} \prod_{j=1}^{i-1} M_t(j)$$

and,

$$R_n^2 = \sigma_{t-n}^2 \prod_{j=1}^{n} M_t(j) = \sigma_{t-n}^2 \prod_{j=1}^{n} (\alpha + \beta \varepsilon_{t-j}^2)$$

As the first step of the proof we show that R_n^2 converges to 0 almost surely (i.e $R_n^2 \xrightarrow{a.s} 0$). Now R_n^2

can be written as,

$$
\begin{aligned}
R_n^2 &= \sigma_{t-n}^2 \exp\left\{\ln\left[\sigma_{t-n}^2 \prod_{j=1}^{n}(\alpha+\beta\varepsilon_{t-j}^2)\right]\right\} \\
&= \sigma_{t-n}^2 \exp\left\{\sum_{j=1}^{n}\ln\left(\alpha+\beta\varepsilon_{t-j}^2\right)\right\} \\
&= \sigma_{t-n}^2 \exp\left\{n\left[\frac{1}{n}\sum_{j=1}^{n}\ln\left(\alpha+\beta\varepsilon_{t-j}^2\right)\right]\right\}
\end{aligned}
$$

Then for fixed t, the strong law of large numbers tells us as $n \to \infty$

$$\frac{1}{n}\sum_{j=1}^{n}\ln\left(\alpha+\beta\varepsilon_{t-j}^2\right) \overset{a.s}{\to} E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right]$$

Thus,

$$R_n^2 \approx \sigma_{t-n}^2 \exp\left\{E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right]\right\}$$

Since σ_{t-n}^2 is bounded, above result tells us $R_n^2 \overset{a.s}{\to} 0$ if and only if $E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right] < 0$.
That implies $\sigma_t^2 = S_n^2 + R_n^2$ with $P\left(R_n^2 \to 0\right) = 1$ if and only if $E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right] < 0$.
Therefore we only need to show that the random variable in equation (1.2) converges if and only if (1.4) holds. Now equation (1.2) can be written as,

$$S^2 = \omega\sum_{i=1}^{\infty}\exp\left\{(i-1)\Big[\frac{1}{i-1}\sum_{j=1}^{i-1}\ln\left(\alpha+\beta\varepsilon_{-j}^2\right)\Big]\right\}$$

Then for fixed t, since equation (1.3) holds, the strong law of large numbers tells us as $i \to \infty$

$$\frac{1}{i-1}\sum_{j=1}^{i-1}\ln\left(\alpha+\beta\varepsilon_{-j}^2\right) \overset{a.s}{\to} E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right]$$

Thus

$$S^2 \approx \omega\sum_{i=1}^{\infty}\exp\left\{E\left[ln\left(\alpha+\beta\varepsilon_0^2\right)\right](i-1)\right\}$$

converges if and only if $\left|\exp\left\{E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right]\right\}\right| < 1$.
That is if and only if $E\left[\ln\left(\alpha+\beta\varepsilon_0^2\right)\right] < 0$, which concludes the proof. Now it needs to be figured

out the conditions that should be satisfied for existence of the vth moment for the GARCH(1,1) model.

Proof of both Theorem 2.1.2 and Theorem 2.2.2 are given together.

Proof of Theorem 2.2.1

If $\nu \geqslant 1$, by Minkowski's inequality we get(by proof of Theorem 2.1, Aue et al. (2006),

$$\begin{aligned} E\left[|X|^{\nu}\right] &\leqslant \left(\sum_{1\leqslant i<\infty}\left[E|\omega\prod_{1\leqslant j\leqslant i-1}(\alpha+\beta\varepsilon_{-j}^{2})|^{\nu}\right]^{1/\nu}\right)^{\nu} \\ &= \omega^{\nu}\left(\sum_{1\leqslant i<\infty}\left[E|\alpha+\beta\varepsilon_{0}^{2}|^{\nu}\right]^{i-1/\nu}\right)^{\nu} < \infty \end{aligned}$$

by assumption 2.5. If $0< \nu <1$. then

$$E\left[|X|^{\nu}\right] leqslant \sum_{1\leqslant i<\infty} E\left[\left|\omega\prod_{1\leqslant j\leqslant i-1}(\alpha+\beta\varepsilon_{-j}^{2})\right|^{\nu}\right] = \omega^{\nu}\sum_{1\leqslant i<\infty}\left[E\left|\alpha+\beta\varepsilon_{0}^{2}\right|^{\nu}\right]^{i-1} < \infty,$$

which completes the proof of the first part of the theorem.

Proof of Theorem 2.1.2 and Theorem 2.2.2

A formula for $E[\sigma_t^2]$ is found by starting from the equation (1.2) derived above, as follows.

$$E[\sigma_t^2] = \omega + \omega E\left[\sum_{i=1}^{\infty} M_t(1)M_t(2)\cdots M_t(i)\right]$$

We can observe that,

$$\begin{aligned} E[M_t(1)M_t(2)\cdots M_t(k)] &= E\{\alpha[M_t(1)\cdots M_t(k-1)]+\beta\varepsilon_{t-k}^{2}[M_t(1)\cdots M_t(k-1)]\} \\ &= \alpha E[M_t(1)\cdots M_t(k-1)]+\beta E[\varepsilon_{t-k}^{2}]E[M_t(1)\cdots M_t(k-1)] \\ &= \alpha E[M_t(1)\cdots M_t(k-1)]+\beta E[M_t(1)\cdots M_t(k-1)] \\ &= (\alpha+\beta)E[M_t(1)\cdots M_t(k-1)] \end{aligned}$$

by solving recursively we get,

$$E[M_t(1)M_t(2)\cdots M_t(k)] = (\alpha+\beta)^{k-1}E[\alpha+\beta\varepsilon_{t-1}^{2}]$$

$$= \quad (\alpha+\beta)^k$$

Thus,

$$\begin{aligned}
E[\sigma_t^2] &= \omega + \omega[(\alpha+\beta) + (\alpha+\beta)^2 + \cdots] \\
&= \omega \sum_{k=0}^{\infty} (\alpha+\beta)^k \qquad\qquad if\ \alpha+\beta<1, \\
&= \frac{\omega}{1-\alpha-\beta} \qquad (3.9)
\end{aligned}$$

Now consider,

$$\begin{aligned}
E\left[x_t\right] &= E\left[\sigma_t \varepsilon_t\right] \\
&= E\left[\sigma_t\right] E\left[\varepsilon_t\right] \qquad \sigma_t^2 \text{ and } \varepsilon_t^2 \text{ are independent} \\
E[x_t] &= 0 \qquad (3.10)
\end{aligned}$$

This also proves that the mean function of x_t is independent of time t, which completes proof of the first part of Theorem 2.1.2 as well.

By using the result obtained in equation (1.2) and $x_t^2 = \sigma_t^2 \varepsilon_t^2$, we get,

$$\begin{aligned}
E[x_t^2] &= E[\sigma_t^2]E[\varepsilon_t^2] \qquad (\sigma_t^2 \text{ is a function of } \varepsilon_t\text{'s and } \varepsilon_t\text{'s are iid}) \\
&= \frac{\omega}{1-\alpha-\beta} \qquad (3.11)
\end{aligned}$$

Then

$$\begin{aligned}
\mathrm{Var}(x_t) &= E[x_t^2] - E[x_t]^2 \\
&= \frac{\omega}{1-\alpha-\beta} \qquad (3.12)
\end{aligned}$$

Also

$$\begin{aligned}
E[x_t x_s] &= E[\sigma_t \varepsilon_t \sigma_s \varepsilon_s] \qquad \text{(where t} \neq \text{s)} \\
&= E[\sigma_t \sigma_s \varepsilon_t \varepsilon_s] \\
&= E[\sigma_t \sigma_s] E[\varepsilon_t \varepsilon_s]
\end{aligned}$$

$$= \quad 0 \tag{3.13}$$

Therfore

$$\text{Cov}(x_t, x_s) = 0$$

Thus the auto-covariance between x_t and x_{t+h} is,

$$\begin{aligned} \gamma_x(h) &= \text{Cov}(x_t, x_{t+h}) \qquad\qquad \text{h} \geq 1 \\ &= E[x_t x_{t+h}] - E[x_t]E[x_{t+h}] \qquad = \quad E[x_t x_{t+h}] \end{aligned}$$

That tells us the auto-covariance function of x_t is also independent of time **t** for each lag **h**, which completes the proof of the Theorem 2.1.2

Therefore the auto-correlation function of x_t can be written as follows.

$$\gamma_x(h) = \begin{cases} \frac{\omega}{1-\alpha-\beta} & h = 0 \\ 0 & \text{otherwise} \end{cases}$$

According to Theorem 2.2.1, if $\text{E}\left[\alpha + \beta\varepsilon_0^2\right]^2$ <1 then $\text{E}\left[\sigma_t^4\right] < \infty$. But,

$$\begin{aligned} E\left[\alpha + \beta\varepsilon_0^2\right]^2 &= E\left[\alpha^2 + \beta^2\varepsilon_0^4 + 2\alpha\beta\varepsilon^2\right] < 1 \\ \text{that is,} \qquad & \alpha^2 + \beta^2 E\left[\varepsilon_0^4\right] + 2\alpha\beta E\left[\varepsilon_0^2\right] < 1 \end{aligned}$$

That gives, if ε_t ~N(0,1) and $\alpha^2 + 3\beta^2 + 2\alpha\beta$ <1 then $\text{E}[x_t^4] < \infty$.

We noted that the ACF of x_t^2 is different from zero. In order to obtain a clear understanding about the fact, it is important to derive the theoretical auto-covariance structure of GARCH(1,1) process.

Proof of Theorem 2.2.3

Starting from equation (2.3) it is possible to show that x_t^2 is equivalent to an ARMA(1,1) process. By subtracting σ_t^2 from x_t^2 we get,

$$x_t^2 - \sigma_t^2 = x_t^2 - \omega - \alpha\sigma_{t-1}^2 - \beta x_{t-1}^2$$

Adding and subtracting the term αx_{t-1}^2 on the right hand side gives,

$$\begin{aligned} x_t^2 - \sigma_t^2 &= x_t^2 - \omega - \alpha\sigma_{t-1}^2 - \alpha x_{t-1}^2 + \alpha x_{t-1}^2 - \beta x_{t-1}^2 \\ x_t^2 - \sigma_t^2 &= x_t^2 - \omega + \alpha(x_{t-1}^2 - \sigma_{t-1}^2) - (\alpha+\beta)x_{t-1}^2 \qquad (3.14) \end{aligned}$$

Let $v_t = x_t^2 - \sigma_t^2$. The expected value and the variance are derived in order to identify the behavior of v_t.

$$\begin{aligned} E[v_t] &= E[x_t^2 - \sigma_t^2] = E[(\varepsilon_t^2 - 1)\sigma_t^2] \\ &= E[(\varepsilon_t^2 - 1)]E[\sigma_t^2] \qquad \text{[since } \varepsilon_t \text{ is independent from } \sigma_t^2\text{]} \\ &= 0 \qquad (3.15) \\ E[v_t v_s] &= E[(x_t^2 - \sigma_t^2)(x_s^2 - \sigma_s^2] \qquad t \neq s \\ &= E[\sigma_t^2\sigma_s^2(\varepsilon_t^2 - 1)(\varepsilon_s^2 - 1)] \\ &= E[\sigma_t^2\sigma_s^2]E[(\varepsilon_t^2 - 1)]E[(\varepsilon_s^2 - 1)] \\ &= 0 \qquad (3.16) \end{aligned}$$

Since both $E[v_t]$ and $\mathrm{Cov}[v_t, v_s]$ are zero, v_t can be regarded as white noise (WN). Then, equation (3.14) can be written down as,

$$x_t^2 = \omega + (\alpha+\beta)x_{t-1}^2 - \alpha v_{t-1} + v_t \qquad (3.17)$$

which is an ARMA(1,1) process and $v_t \sim WN(0, \sigma_v^2)$.

By subsequent substitutions, an ARMA(1,1) process can expressed as an infinite order MA process as follows.

$$
\begin{aligned}
x_t^2 &= \omega + (\alpha+\beta)\left[\omega + (\alpha+\beta)x_{t-2}^2 - \alpha v_{t-2} + v_{t-1}\right] - \alpha v_{t-1} + vt \\
&= \omega + \omega(\alpha+\beta) + (\alpha+\beta)^2 x_{t-2}^2 - \alpha(\alpha+\beta)v_{t-2} + \left[(\alpha+\beta)-\alpha\right]v_{t-1} + v_t \\
&\vdots \\
&= \omega\left[1 + (\alpha+\beta) + (\alpha+\beta)^2 + \cdots\right] + v_t + \beta\big[v_{t-1} + (\alpha+\beta)v_{t-2} + \\
&\quad (\alpha+\beta)^2 v_{t-3} + \cdots\big] \qquad (3.18)
\end{aligned}
$$

Let (2.5) hold with $\nu \geqslant 2$. That implies $E[\varepsilon_t^4]$ is finite and thus $\mathrm{Var}[v_t]$ exists. Let $\mathrm{Var}[v_t] = \sigma_v^2$. Therefore,

$$
\begin{aligned}
\gamma_{x^2}(0) = \mathrm{Var}[x_t^2] &= \sigma_v^2 + \beta^2\sigma_v^2\left[1 + (\alpha+\beta)^2 + (\alpha+\beta)^4 + \cdots\right] \\
&= \sigma_v^2\left[1 + \frac{\beta^2}{1-(\alpha+\beta)^2}\right] \qquad \text{when} \quad \alpha+\beta < 1 \qquad (3.19)
\end{aligned}
$$

Since $v_t \sim WN(0, \sigma_v^2)$, we know that,

$$
\begin{aligned}
E[x_t^2] &= \omega\left[1 + (\alpha+\beta) + (\alpha+\beta)^2 + \cdots\right] \\
&= \frac{\omega}{1-(\alpha+\beta)} \qquad \text{when} \quad \alpha+\beta < 1
\end{aligned}
$$

It can also be shown that,

$$
\begin{aligned}
x_t^2 x_{t-1}^2 = &\left(\frac{\omega}{1-(\alpha+\beta)} + v_t + \beta v_{t-1} + \beta(\alpha+\beta)v_{t-2} + \cdots + \beta(\alpha+\beta)^{h-1} v_{t-h} + \cdots\right) \\
&\left(\frac{\omega}{1-(\alpha+\beta)} + v_{t-1} + \beta v_{t-2} + \cdots + \beta(\alpha+\beta)^{h-2} v_{t-h} + \beta(\alpha+\beta)^{h-1} v_{t-h-1} + \cdots\right)
\end{aligned}
$$

Then by taking the expectation of both sides, we get,

$$
\begin{aligned}
E[x_t^2 x_{t-1}^2] &= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \sigma_v^2\left(\beta + \beta^2(\alpha+\beta) + \beta^2(\alpha+\beta)^3 + \cdots\right) \\
&= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \beta\sigma_v^2 + \beta^2(\alpha+\beta)\sigma_v^2\left(1 + (\alpha+\beta)^2 + (\alpha+\beta)^4 + \cdots\right)
\end{aligned}
$$

$$= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \beta\left[1+\frac{\beta(\alpha+\beta)}{1-(\alpha+\beta)^2}\right]\sigma_v^2$$

Then

$$\begin{aligned}\gamma_{x^2}(1) &= \operatorname{Cov}\left(x_t^2 x_{t-1}^2\right) = E[x_t^2 x_{t-1}^2] - E[x_t^2]E[x_{t-1}^2] \\ &= \beta\left[1+\frac{\beta(\alpha+\beta)}{1-(\alpha+\beta)^2}\right]\sigma_v^2\end{aligned}$$

Similarly,

$$\begin{aligned}E[x_t^2 x^2 t-h] &= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \beta(\alpha+\beta)^{h-1}\sigma_v^2 + \sigma_v^2\left(\beta^2(\alpha+\beta)^h + \beta^2(\alpha+\beta)^{h+2} + \cdots\right) \\ &= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \beta(\alpha+\beta)^{h-1}\sigma_v^2 + \beta^2(\alpha+\beta)^h\left[\frac{1}{1-(\alpha+\beta)^2}\right]\sigma_v^2 \\ &= \left(\frac{\omega}{1-(\alpha+\beta)}\right)^2 + \beta(\alpha+\beta)^{h-1}\left[1+\frac{\beta(\alpha+\beta)}{1-(\alpha+\beta)^2}\right]\sigma_v^2\end{aligned}$$

Then

$$\begin{aligned}\gamma_{x^2}(h) &= \operatorname{Cov}\left(x_t^2 x_{t-h}^2\right) = E[x_t^2 x_{t-h}^2] - E[x_t^2]E[x_{t-h}^2] \\ &= \beta(\alpha+\beta)^{h-1}\left[1+\frac{\beta(\alpha+\beta)}{1-(\alpha+\beta)^2}\right]\sigma_v^2 \\ &= (\alpha+\beta)^{h-1}\gamma_{x^2}(1)\end{aligned}$$

Therefore the ACF of x_t^2 can be represented by,

$$\rho_{x^2}(h) = \begin{cases} 1 & \text{h} = 0 \\ \beta\left(\frac{1-\alpha(\alpha+\beta)}{1-\alpha(\alpha+2\beta)}\right) & \text{h} = 1 \\ (\alpha+\beta)^{h-1}\rho_{x^2}(1) & \text{h}>1 \end{cases}$$

Thus it is clear that the ACF of x_t^2 is different from zero and decays exponentially.

Chapter 3

The Statistical Analysis of GARCH(1,1) Processes

3.1 Maximum likelihood estimation

The Maximum Likelihood Estimates(MLE) can be obtained only if the density of the data are known. Assuming the density of ε_k is known to be $h(x)$, the conditional likelihood function of the observations $x_1, x_2, \ldots, x_n$ of the GARCH(1,1) process when x_0 and σ_0^2 are given, can be written as,

$$L_n(x, \theta) = \prod_{1 \leqslant k \leqslant n} \frac{1}{\sigma_k(\theta)} h\left(\frac{x_k}{\sigma_k(\theta)}\right) \qquad \text{where} \qquad \theta = (\omega, \alpha, \beta)$$

Since σ_k^2 also depends on parameters (ω, α, β), it is needed to solve for σ_k^2 before maximizing the likelihood function. When σ_0^2 and x_0 are given, the solution for σ_k^2 can be written in the following form.

$$\sigma_k^2 = \omega \sum_{i=0}^{k-1} \alpha^i + \alpha^k \sigma_0^2 + \beta \sum_{i=0}^{k-1} \alpha^i x_{k-1-i}^2$$

Hence the parameters can be estimated after substituting σ_k^2 to the likelihood function and maximizing it respect to GARCH(1,1) parameters. These estimated parameters are known as conditional likelihood parameters of the GARCH(1,1) process.

3.2 Testing for parameter consistency

Detecting possible change points is a very important task not only for GARCH (1,1) model but also for any kind of time series model. Since all the inferences are made based on estimated parameters, if there exists any changes in these parameters those inferences will not be valid any longer.
In order to figure out the structural changes (change points), the test proposed in Theorem 2.1 in Berkes et al. (2004) is rewritten for a GARCH (1,1) model. In this study, we assume that innovations of GARCH (1,1) model are normally distributed with mean zero and variance one, that is $\epsilon_t \sim N(0,1)$. Under null hypothesis, by assuming no change point exists, we can write down the theorem as follows.

Theorem 3.2.1 *Suppose that $\{x_t : t \in \mathbb{Z}\}$ is a GARCH(1,1) process specified by the equations (2.2), (2.3) and satisfies the equation (1.4). Under a set of assumptions (see Berkes et al. (2004) for more details), then.*

$$\frac{1}{n}\left[\sum_{1\leqslant i\leqslant nt} \tilde{l_i'}\left(\hat{\theta}_n\right) - \frac{[nt]}{n}\sum_{1\leqslant i\leqslant n} \tilde{l_i'}\left(\hat{\theta}_n\right)\right] \qquad D^{-1}\left[\sum_{1\leqslant i\leqslant nt} \tilde{l_i'}\left(\hat{\theta}_n\right) - \frac{[nt]}{n}\sum_{1\leqslant i\leqslant n} \tilde{l_i'}\left(\hat{\theta}_n\right)\right]^T \rightarrow \sum_{1\leqslant i\leqslant 3} B_i^2\left(t\right) inD\left[0,1\right], \tag{2.1}$$

where

$$D = \mathrm{Cov}\left[l_0'\left(\theta\right)\right] = -E\left[l_0''\left(\theta\right)\right]$$

and $\{B_i\left(t\right), 0 \leqslant t \leqslant 1\}, i = 1, \cdots, 3,$ *are independent Brownian bridges.*

where,

$$\begin{aligned} \theta &= (\omega, \alpha, \beta) \\ l_k\left(\epsilon_k, \theta\right) &= -\frac{1}{2}ln\left(2\pi\right) - \frac{1}{2}ln\left(\sigma_k^2\right) - \frac{1}{2}\frac{x_k^2}{\sigma_k^2} \end{aligned} \tag{2.2}$$

By assuming σ_0 and $x_k; k = 0, 1, 2\cdots, n$ are known, we can write the solution for σ_k^2 by,

$$\sigma_k^2 = \omega\sum_{i=0}^{k-1}\alpha^i + \alpha^k\sigma_0^2 + \beta\sum_{i=0} k - 1\alpha^i x_{k-1-i}^2 \tag{2.3}$$

The matrix $\mathbf{D}$ can be estimated by $n^{-1}\sum_{1\leqslant i\leqslant n}\left[\tilde{l}_i'\left(\hat{\theta}_n\right)\right]^T\left[\tilde{l}_i'\left(\hat{\theta}_n\right)\right]$ or by $-n^{-1}\sum_{1\leqslant i\leqslant n}\left[\tilde{l}_i''\left(\hat{\theta}_n\right)\right]$. A functional of the above expression is proposed as the statistic. The critical values for $\int_0^1 U_3(t)dt = \int_0^1\sum_{i=1}^3 B_i^2(t)dt$ are in Table 4 on page 444 of Kiefer (1959).

We also know that,

$$l_k'(\theta) = \left[\frac{\partial l_k}{\partial \omega}, \frac{\partial l_k}{\partial \alpha}, \frac{\partial l_k}{\partial \beta}\right] = \left[l_{k,1}'(\theta), l_{k,2}'(\theta), l_{k,3}'(\theta)\right]$$

and denote,

$$l_k''(\theta) = \begin{bmatrix} \frac{\partial l_{k,1}'(\theta)}{\partial \omega} & \frac{\partial l_{k,2}'(\theta)}{\partial \omega} & \frac{\partial l_{k,3}'(\theta)}{\partial \omega} \\ \frac{\partial l_{k,1}'(\theta)}{\partial \alpha} & \frac{\partial l_{k,2}'(\theta)}{\partial \alpha} & \frac{\partial l_{k,3}'(\theta)}{\partial \alpha} \\ \frac{\partial l_{k,1}'(\theta)}{\partial \beta} & \frac{\partial l_{k,2}'(\theta)}{\partial \beta} & \frac{\partial l_{k,3}'(\theta)}{\partial \beta} \end{bmatrix}$$

Now we derive expressions for derivative terms.

$$\begin{aligned} \frac{\partial l_k}{\partial \omega} &= -\frac{1}{2\sigma_k}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i\right] + \frac{1}{2\sigma_k^4}\left[x_k^2\sum_{0\leqslant i\leqslant k-1}\alpha^i\right] \\ &= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i\left[x_k^2-\sigma_k^2\right]\right] \end{aligned}$$

where,

$$\sigma_k^2 = \omega\sum_{0\leqslant i\leqslant k-1}\alpha^i + \alpha^k\sigma_0^2 + \beta\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2$$

$$\begin{aligned} \frac{\partial l_k}{\partial \alpha} &= -\frac{1}{2\sigma_k^2}\left[\omega\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i + (k-1)\alpha^{k-2}\sigma_0^2 + \beta\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2\right] + \\ &\quad \frac{1}{2\sigma_k^4}\left[x_k^2\left[\omega\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i + (k-1)\alpha^{k-2}\sigma_0^2 + \beta\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2\right]\right] \\ &= \frac{1}{2\sigma_k^4}\left[B_{k-2}\left[x_k^2-\sigma_k^2\right]\right] \end{aligned}$$

where,

$$B_{k-2} = \left[\omega\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i + k\alpha^{k-1}\sigma_0^2 + \beta\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2\right]$$

$$
\begin{aligned}
\frac{\partial l_k}{\partial \beta} &= -\frac{1}{2\sigma_k^2}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right] + \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right] \\
&= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\left[x_k^2-\sigma_k^2\right]\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \omega^2} &= \frac{1}{2\sigma_k^4}\left[\left(\sum_{0\leqslant i\leqslant k-1}\alpha^i\right)^2\right] - \frac{1}{2\sigma_k^6}\left[2x_k^2\left(\sum_{0\leqslant i\leqslant k-1}\alpha^i\right)^2\right] \\
&= \frac{1}{2\sigma_k^6}\left[\left(\sum_{0\leqslant i\leqslant k-1}\alpha^i\right)^2\left[\sigma_k^2-2x_k^2\right]\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \alpha \partial \omega} &= -\frac{1}{2\sigma_k^4}\left[\sigma_k^2\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i - \sum_{0\leqslant i\leqslant k-1}\alpha^i B_{k-2}\right] + \\
&\quad \frac{1}{2\sigma_k^8}\left[x_k^2\left[\sigma_k^4\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i - \sum_{0\leqslant i\leqslant k-1}\alpha^i 2\sigma_k^2 B_{k-2}\right]\right] \\
&= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i\left[x_k^2-\sigma_k^2\right]\right] + \frac{1}{2\sigma_k^6}\left[B_{k-2}\sum_{0\leqslant i\leqslant k-1}\alpha^i\left[\sigma_k^2-2x_k^2\right]\right] \\
&= \frac{\partial^2 l_k}{\partial \omega \partial \alpha}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \beta \partial \omega} &= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right] - \frac{1}{2\sigma_k^6}\left[2x_k^2\sum_{0\leqslant i\leqslant k-1}\alpha^i\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right] \\
&= \frac{1}{2\sigma_k^6}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\left[\sigma_k^2-2x_k^2\right]\right] \\
&= \frac{\partial^2 l_k}{\partial \omega \partial \beta}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \alpha^2} &= -\frac{1}{2\sigma_k^4}\Big[\sigma_k^2[\omega \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i + (k-1)(k-2)\alpha^{k-3}\sigma_0^2 + \\
&\quad \beta \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i x_{k-3-i}^2] - B_{k-2}^2\Big] + \\
&\quad \frac{1}{2\sigma_k^8}\Big[x_k^2[\sigma_k^4[\omega \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i + (k-1)(k-2)\alpha^{k-3}\sigma_0^2 + \\
&\quad \beta \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i x_{k-3-i}^2] - 2B_{k-2}^2\sigma_k^2]\Big] \\
&= \frac{1}{2\sigma_k^4}\left[C_{k-3}\left[x_k^2 - \sigma_k^2\right]\right] + \frac{1}{2\sigma_k^6}\left[B_{k-2}^2\left[\sigma_k^2 - 2x_k^2\right]\right]
\end{aligned}
$$

where,

$$
C_{k-3} = \left[\omega \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i + k(k-1)\alpha^{k-2}\sigma_0^2 + \beta \sum_{0\leqslant i\leqslant k-3}(i+1)(i+2)\alpha^i x_{k-3-i}^2\right]
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \beta \partial \alpha} &= -\frac{1}{2\sigma_k^4}\left[\sigma_k^2 \sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2 - B_{k-2}\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right] + \\
&\quad \frac{1}{2\sigma_k^8}\left[x_k^2\left[\sigma_k^4 \sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2 - 2B_{k-2}\sigma_k^2 \sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right]\right] \\
&= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-2}(i+1)\alpha^i x_{k-2-i}^2\left[x_k^2 - \sigma_k^2\right]\right] + \\
&\quad \frac{1}{2\sigma_k^6}\left[B_{k-2}\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\left[\sigma_k^2 - 2x_k^2\right]\right] \\
&= \frac{\partial^2 l_k}{\partial \alpha \partial \beta}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_k}{\partial \beta^2} &= \frac{1}{2\sigma_k^4}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right]^2 - \frac{1}{2\sigma_k^6}\left[2x_k^2\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right]^2\right] \\
&= \frac{1}{2\sigma_k^6}\left[\sum_{0\leqslant i\leqslant k-1}\alpha^i x_{k-1-i}^2\right]^2\left[\sigma_k^2 - 2x_k^2\right]
\end{aligned}
$$

3.3 Bootstrapping a GARCH(1,1) process

Bootstrapping is a technique of regenerating data based on an available sample. Since this procedure involves repeating the original procedure with many simulated sets of data, these are sometimes called "computer intensive methods". Bootstrapping can be categorized into two based on the fact that the distribution of the data is known and unknown. When the distribution is known, it is known as parametric bootstrapping methods and otherwise it is known as non-parametric bootstrapping methods.

a) Parametric bootstrapping

   Suppose the distribution of the available data $y_1, y_2, \ldots, y_n$ is known and the cumulative distribution function (CDF) and probability density function (PDF) are denoted by $F_\theta(y)$ and $f_\theta(y)$ respectively. When θ is estimated by $\hat{\theta}$, often its maximum likelihood estimate; its substitution in the distribution function gives the fitted distribution, with $\hat{F}(y) = F_{\hat{\theta}}(y)$, which can be used to generate data.

b) Non-parametric bootstrapping

   Since the distribution function of available sample of data is unknown, the estimated distribution function is determined by putting equal probabilities on the data values $y_1, y_2, \ldots, y_n$. Then the re-sampling can be done by drawing a random sample with replacement from the data.

Bootstrapping a GARCH(1,1) process is not too easy since the correlation structure of both data and the volatility process needs to be taken into account. Therefore the following procedure is implemented for bootstrapping.

1. Parameter estimates $\hat{\omega}, \hat{\alpha}$ and $\hat{\beta}$ are obtained by fitting the GARCH(1,1) process to the data.
2. The volatility σ_t^2 is estimated by $\hat{\sigma}_t^2 = \hat{\omega} + \hat{\alpha}\hat{\sigma}_{t-1}^2 + \hat{\beta}x_{t-1}^2$, where $\hat{\sigma_0^2} = \frac{\hat{\omega}}{1-\hat{\alpha}-\hat{\beta}}$ and $x_0 = \sqrt{\frac{\hat{\omega}}{1-\hat{\alpha}-\hat{\beta}}}$.
3. The error series is estimated by, $\hat{\varepsilon}_t = \frac{x_t}{\hat{\sigma}_t}$.

4. Under the assumption that the error terms follow an iid $N(0,1)$ distribution, the error distribution is estimated as; $\varepsilon \sim N\left(mean(\hat{\varepsilon}), \mathrm{Var}(\hat{\varepsilon})\right)$.

5. Then bootstrapped data and volatility terms are generated as follows.

   - $\sigma^2_{t(boot)} = \hat{\omega} + \hat{\alpha}\hat{\sigma}^2_{t-1(boot)} + \hat{\beta}x^2_{t-1(boot)}$, where $\sigma^2_{0(boot)} = \frac{\hat{\omega}}{1-\hat{\alpha}-\hat{\beta}}$ and $x_{0(boot)} = \sqrt{\frac{\hat{\omega}}{1-\hat{\alpha}-\hat{\beta}}}$.
   - $x_{t(boot)} = \sigma_{t(boot)}\varepsilon^*_t$, where ε^*_t is randomly generated from the estimated distribution of ε_t.

In Chapter 4 we will discuss how these described theories and methodologies work on a real world application.

Chapter 4

An Application to Dow Jones Index Data

The suitability of GARCH(1,1) model for the returns of Dow Jones Index (DWJ) is compared with the downloaded actual data values from Jan 1997 through Dec 2006 and simulated data according to the fitted GARCH(1,1) model. This section also focuses on detecting change points in parameters of the log-return series by applying the test proposed by Berkes et al. (2004) and assessing the power of the test.

4.1 Fitting a GARCH(1,1) model to data

Before fitting the GARCH(1,1) model the log difference transformation given in (0.1) is applied in order to capture the common features discussed before. The GARCH(1,1) model, which fits the log-return data is,

$$x_t^2 = \omega + \alpha\sigma_{t-1}^2 + \beta x_{t-1}^2$$

where

$$\begin{aligned} &\omega = 9.77493 \times 10^{-7} \qquad \alpha = 0.91031 \\ &\beta = 0.08516 \end{aligned} \tag{1.1}$$

Since the coefficients $\alpha = 0.91031$ and $\beta = 0.08516$, and $\alpha + \beta = 0.99547 < 1$, it is clear that the fitted GARCH(1,1) process is both strictly and weakly stationary. However it is noticeable that under the normal assumption of innovation distribution $E[|\alpha + \beta\varepsilon_0^2|^2] = 1.0055$, which is slightly greater than one. That gives evidence for the infinite fourth moment of log-returns. But still we can apply the the test proposed by Berkes, Horváth and Kokoszka (BHK test) since the only condition that the series must satisfy is the $E[\varepsilon_0^4] < \infty$, and the estimated value; $E[\hat{\varepsilon}_0^4] = 4.43077$, which is finite under the normal assumptions.
In order to assess how well the fitted GARCH(1,1) process captures the features of the log-return series, another set of data with the same number of observations as the original series has, is simulated according to the model equation given in (1.1). In the following section we discuss the features of DWJ index data and the corresponding simulated series to asses the compatibility of the two series.

4.2 Time plots

We start the assessment of the behavior of DWJ and simulated data by focusing on the corresponding time plots over the time period of ten years (Figure 4.1). As a whole, it is apparent that the log-returns always fluctuate about zero and the mean value confirms that suggestion by standing at 2.623×10^{-4}. Even though the log-returns oscillates vigorously until about mid 2003, afterward it shows lesser volatility. Therefore the time plot can roughly be divided in to two periods one is from January 1997 through about 2003 and the other period is from 2003 through December 2006. These two different behaviors motivate to investigate for structural changes in GARCH(1,1) model fitted for the full data series. In addition to its highly volatile behavior it can be clearly seen few significant shocks affected on the first half of the process. One of the largest shocks has been detected on the very next opening day after September 11th attack in 2001, which has the value of -0.07396. It is also detected that the largest drop in US stock market has occurred on October 27, 2007, which is the opening day after closing the market for the respective weekend before 27th. However a significant reason for this huge shock could not be identified. For the simulated log-returns 4.1, though the mean is almost zero (-8.314×10^{-04}), the appearance of the time plot is not same as the innovative one.It shows vigorous ups and downs until about 2002, then the fluctuation becomes less compared with the period before 2002. Again it shows some significant fluctuations after 2006.

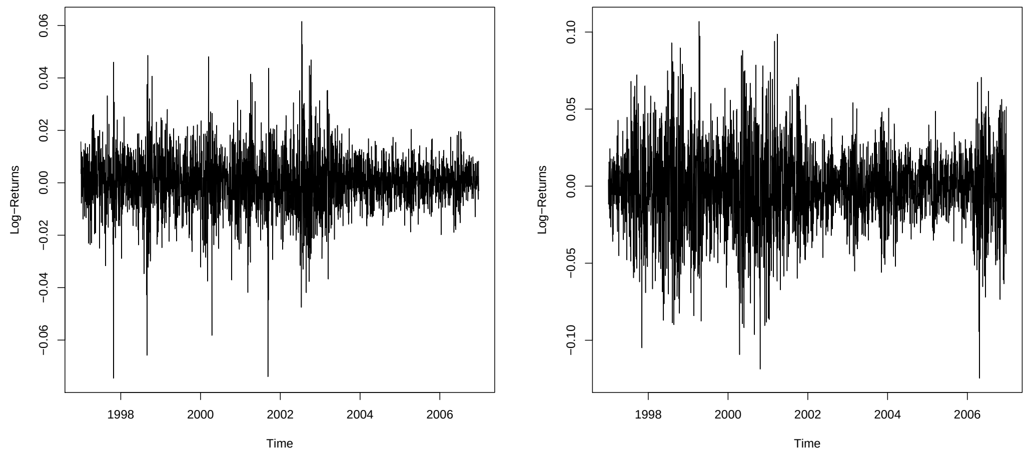

Figure 4.1: *Comparison of daily log-returns of Dow Jones Index values (left) and simulated log-returns (right)*

4.3 The auto-correlation functions for the log-returns

We proved in Theorem 2.2.2 that the theoretical ACF of a GARCH(1,1) model is zero after the lag zero. It is observable that most of the spikes after the first lag are insignificant for log-returns of DWJ data (Figure 4.2) and significant changes cannot be seen in the ACF corresponding to the simulated series.

4.4 The auto-correlation function for the squares of log-returns

We expect exponentially decaying ACF for squares of log-returns according to the result obtained under the Theorem 2.2.3. Figure 4.3 illustrates a slow decaying pattern for squares of log-returns and few more unusual spikes at larger lags as well. When it compares with the ACF of squares of simulated log-returns, very slow exponential decay is visible and spikes remains significant for higher number of lags than in the ACF for the square of DWJ log-returns.

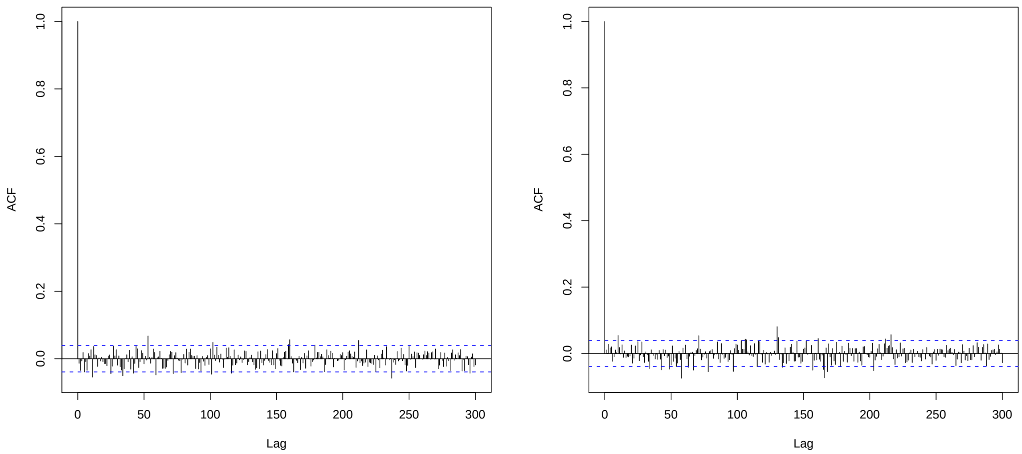

Figure 4.2: *Comparison of ACFs of daily log-returns of Dow Jones Index values (left) and simulated log-returns (right)*

4.5 Detection of change points

In this section we discuss the results obtained by applying the test proposed by Berkes, Horváth and Kokoszka as mentioned before. The critical value was picked based on the area under the curve; that is the value of $\int_0^1 U_3(t)dt$ where $U_3(t) = \sum_{1 \leqslant i \leqslant 3} B_i^2(t)$. The corresponding critical values are given in Table 4 on page 444 of Kiefer(1959).

Since the critical values are not available for the maximum of the limiting distribution, bootstrapping technique is applied as an alternative procedure. Under the bootstrapping method, the maximum value of the test statistic is compared with bootstrapped sample maximum test statistic values.

The following Figure 4.4 shows how the test statistic values varies over t, $0 \leqslant t \leqslant 1$.

It is found that the mean value of all the test statistic values, which is the area under the step function, is 2.16247 and is much higher than the corresponding critical value at 5% significance level, 1.00018 (Table 4, Kiefer, 1959). That provides evidence for the presence of a change point.

It is also detected that the maximum test statistic value, 5.911487 is achieved at the 1480th value and the p-value corresponding of 10,000 bootstrapped samples is computed as 0.0067. Since, according to both procedures, the test statistic value is highly significant, the most plausible point to have a change point is detected as the 1480th value of the series. Thus the series is divided at 1480th data

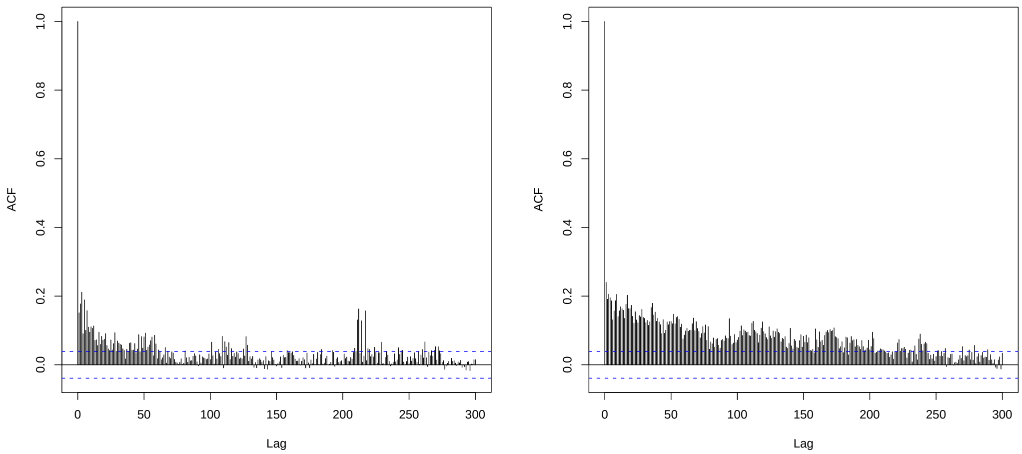

Figure 4.3: *Comparison of ACFs of squared log-returns of Dow Jones Index values (left) and simulated squared log-returns (right)*

value into two sub-samples and applied the same procedures to check for any other change point presence in those subsequent data sets. In order to accomplish that task, separate GARCH(1,1) models are fitted for two sub-samples and test statistic values are obtained. The corresponding two GARCH(1,1) models can be given as follows.

Model 1: (From January 2nd, 1997 through November 10th, 2002)

$$x_t^2 = \omega_1 + \alpha_1 \sigma_{t-1}^2 + \beta_1 x_{t-1}^2; \qquad 1 \leqslant t \leqslant 1480.$$

where,

$$\omega_1 = 8.3132 \times 10^{-6} \qquad \alpha_1 = 0.8464$$
$$\beta_1 = 0.10597$$

Model 2: (From November 11th, 2002 through December 29th, 2006)

$$x_t^2 = \omega_2 + \alpha_2 \sigma_{t-1}^2 + \beta_2 x_{t-1}^2; \qquad 1481 \leqslant t \leqslant 2515.$$

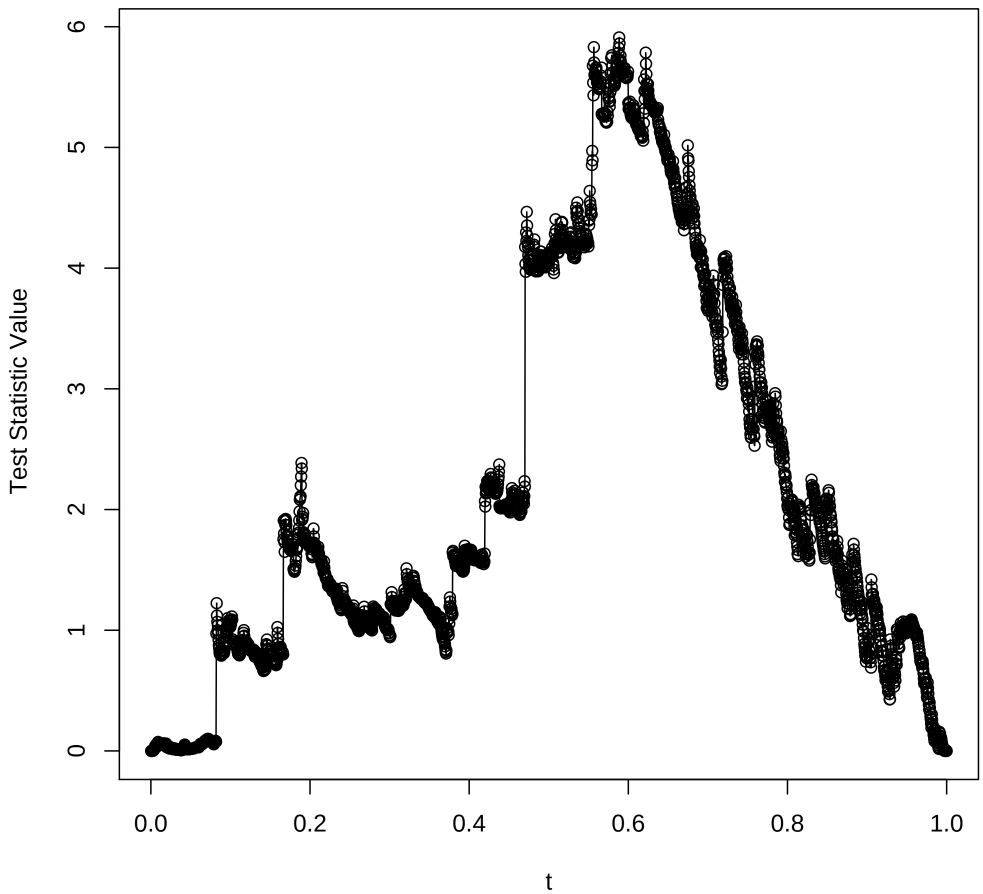

Figure 4.4: *Test statistic values for Log-returns of Dow Jones Index series*

where

$$\omega_1 = 7.1816 \times 10^{-7} \qquad \alpha_1 = 0.9354$$
$$\beta_1 = 0.0506$$

Clearly two mean test statistic values Table 4.1 are less than the critical value of 1.00018. Thus we can conclude that there exist no change points in the two sub-samples at 5% significance level. These conclusions are confirmed by bootstrap results since both p-values are greater than 0.05. The following plots (Figure 4.5) illustrate the how the test statistic value changes over the time and the

Table 4.1: Sub-sample test results

Sub-sample	Mean Test Statistic Value	Max. of Test Statistic Value	Bootstrap P-Value
Sample 1	0.6574	1.64256	0.3780
Sample 2	0.9572	3.16177	0.0695

place where it achieves the maximum value.

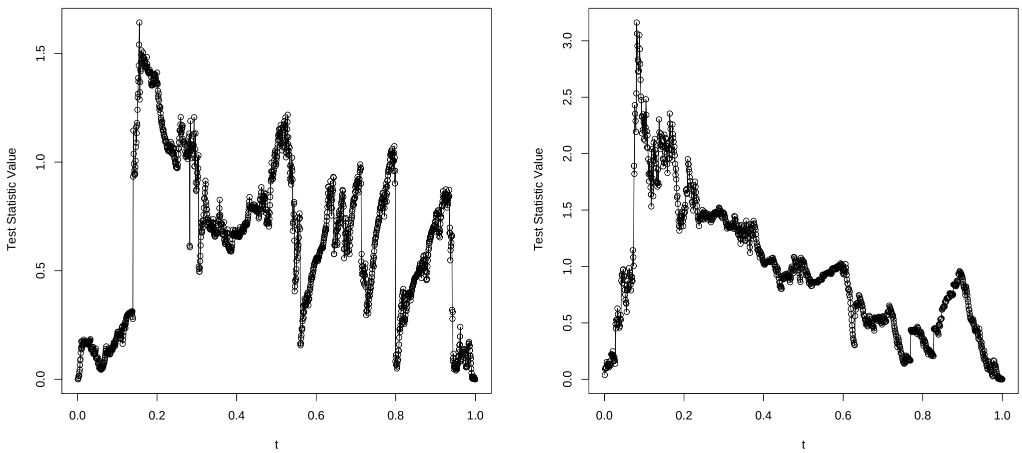

Figure 4.5: *Test Statistic Values for the first sub-sample (left) and the second sub-sample (right)*

Therefore finally we can conclude that there exists only one change point in log-returns of Dow Jones Index series and it exists at 1480th time point. That is, the structure of the GARCH(1,1) series changes after November 20th, 2002.

4.6 Type one error and the power of the test

In this section, we assess the type one error and the power of the change point test. The critical value (1.00018) obtained from Kiefer, (1959) is used as the cutoff point for both tasks.

In order to assess the type one error, 10,000 samples of the same size as the original data set with the fitted GARCH(1,1) parameter values are simulated under the assumption of normal(0,1). It is found out that the type one error of the test is 0.0496. Since the type one error, which is the probability of rejecting the null hypothesis when it is true, is less than 0.05, we can conclude the test proposed by Berkes, Horváth, and Kokoszka is a good test to identify change points in terms of type one error.

Assessing the power of the test is also very crucial in finding a good test. Following procedure is followed to determine the power of the test.

1. Two separate GARCH(1,1) models are fitted to two sub-samples.

2. Two separate samples of sizes corresponding to sizes of two sub-samples are generated with respective GARCH(1,1) fitted model parameters in (1) under the normal(0,1) assumption.

3. Two generated sub-samples are combined as it makes a single sample of size corresponding to the size of the original data set and fitted one GARCH(1,1) model.

4. The mean of the test statistic is derived for the simulated sample in (3).

5. Same procedure is repeated for 10000 times.

It is found that the power of the test, for the full data series is 62.84%. Since the calculated test statistic value for the second sub-sample of the first split is slightly insignificant, the power of the test on this sample is also checked before makinf a conclusion about its power.

In order to assess the power, the second sub-sample is divided into two more samples at the point where it showed the highest test statistic value, which is at the 1564th data point. Again the same procedure that is used to check the power, is applied and derived the mean test statistic values for 10,000 simulations. It is found that 91.05 percent of the simulated values are identified as insignificant, even though the samples are generated under the alternative hypothesis that is by assuming there exist a change point at the 1564th observation. Thus the BHK test could not identify the point that we divide the sub-sample into two parts as a change point quiet well. This shows that though the power of the test on the full data series is quiet good, its power on the second sub-sample is questionable.

4.7 Conclusion and Future Work

We identified exactly one change point in Dow Jones Stock Return series from January 1997 through December 2006. It is concluded that there exist two difference GARCH(1,1) models, one is for the series until November 20th 2002 and the other one is for the series starting from November

20th 2002 and ending December 29th 2006. The above results also show that the test proposed by Berkes, Horváth, and Kokoszka works well on log-returns of Dow Jones index series in terms of type one error. But it is not easy to make a clear conclusion regarding the power of the test. Even though it shows quite good power for the full data set, its power on the second sub-sample of the first split is quiet poor.

Still it is remaining to do more power simulations before making a clear decision about the power of the test. This application study only focused on error series, which is normally distributed with mean zero and variance one, it is needed to check how well this test works on other error distributions such as t-distribution. After completing all these tasks we can compare the performance of this test with number of other tests proposed to identify the structural changes in GARCH(1,1) model.

Appendix

R-Code for the first order partial derivative vector and second order partial derivative matrix.

```
Deriv = function(a, b, c, s, vec, k, mA, mB, mC){
A = a*cumsum(b^(0:(k-1))) + (b^(1:k)*s) + c*mA
B = a*cumsum((0:(k-1))*c(0,b^(0:(k-2)))) + ((1:k)*b^(0:(k-1))*s) + c*mB
C = a*cumsum((0:(k-1))*c(0,(0:(k-2)))*c(0,0,b^(0:(k-3)))) + ((1:k)*(0:(k-1))*b^(-1:(k-2))*s) + c*mC

dl_omega = cumsum(cumsum(b^(0:(k-1)))*(vec[2:length(vec)]^2 - A)/(2*A^2))
dl_alpha = cumsum(B*(vec[2:length(vec)]^2 - A)/(2*A^2))
dl_beta = cumsum(mA*(vec[2:length(vec)]^2 - A)/(2*A^2))

d2l_omega = cumsum(b^(0:(k-1)))^2*(A - (2*vec[2:length(vec)]^2))/(2*A^3)
d2l_alpha.omega = (cumsum((0:(k-1))*c(0,b^(0:(k-2))))*(vec[2:length(vec)]^2 - A)/(2*A^2)) +
(B*cumsum(b^(0:(k-1)))*(A - (2*vec[2:length(vec)]^2))/(2*A^3))
d2l_beta.omega = cumsum(b^(0:(k-1)))*mA*(A - (2*vec[2:length(vec)]^2))/(2*A^3)
d2l_alpha = (C*(vec[2:length(vec)]^2 - A)/(2*A^2)) + (B^2*(A - (2*vec[2:length(vec)]^2))/(2*A^3))
d2l_beta.alpha = (mB*(vec[2:length(vec)]^2 - A)/(2*A^2)) + (B*mA*(A - (2*vec[2:length(vec)]^2))/(2*A^3))
d2l_beta = mA^2*(A - (2*vec[2:length(vec)]^2))/(2*A^3)

d1.vec = cbind(dl_omega, dl_beta, dl_alpha)
d2.matrix = matrix(c(sum(d2l_omega), sum(d2l_beta.omega), sum(d2l_alpha.omega),sum(d2l_beta.omega),
sum(d2l_beta), sum(d2l_beta.alpha), sum(d2l_alpha.omega),sum(d2l_beta.alpha), sum(d2l_alpha)), nrow=3, ncol=3)
return(list(d1.vec=d1.vec, d2.matrix=d2.matrix))
}
```

R-Code for the Horváth, Berkes and Kokozska test function.

```
Test = function(index=1, vec, matrix){
return((1/length(vec))*t(vec[[index]]-(index/length(vec))*vec[[length(vec)]])%*%solve(-matrix/length(vec))
 %*%(vec[[index]]-(index/length(vec))*vec[[length(vec)]]))
}
```

R-Code for the Change-point test (Full-Data Set)

```
data <- scan("data.txt")

dwj = 0
dwj = data[length(data):1]
dwj.log <- log(dwj)
dwj.log.diff <- diff(dwj.log)

garch11 = garchFit(~garch(1,1), dwj.log.diff)
coef <- garch11@fit$coef[2:4]

x0 = sqrt(coef[1]/(1-coef[2]-coef[3]))
sigma0_2 = (coef[1]/(1-coef[2]-coef[3]))

newx = c(x0, dwj.log.diff)

n = length(dwj.log.diff)-1
m1<-matrix(rep(c(coef[3]^(0:n),0),n+1), nrow=n+1, ncol=n+1)
m1[upper.tri(m1)]=0

mA=m1%*%(newx[1:length(dwj.log.diff)]^2)

m21<-matrix(rep(c(0, coef[3]^(0:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m22<-matrix(rep(c(0, (1:n),0),n+1), nrow=n+1, ncol=n+1)
m2<-m21*m22

mB=m2%*%(newx[1:length(dwj.log.diff)]^2)

m31<-matrix(rep(c(0,0, coef[3]^(0:(n-2)),0),n+1), nrow=n+1, ncol=n+1)
m32<-matrix(rep(c(0,0,(1:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m33<-matrix(rep(c(0,0,(2:n),0),n+1), nrow=n+1, ncol=n+1)
m3<-m31*m32*m33

mC=m3%*%(newx[1:length(dwj.log.diff)]^2)

d11.vec = Deriv(coef[1], coef[3], coef[2], sigma0_2, newx, length(dwj.log.diff), mA, mB, mC)$d1.vec
d.matrix = Deriv(coef[1], coef[3], coef[2], sigma0_2, newx, length(dwj.log.diff), mA, mB, mC)$d2.matrix

d.vec=0
for(i in 1:length(d11.vec[,1])){
d.vec[i]=list(d11.vec[i,])
}

T.test = sapply(1:length(d.vec), Test, d.vec)
T.test
```

```
max(T.test)
mean(T.test)

x_axis = function(index=1, vec){
return(index/(length(vec)))
}

xvec=c(rep(1, times=length(d.vec)))
x.value = sapply(1:length(xvec), x_axis, xvec)

plot(x.value, T.test, xlab="t", ylab="Test Statistic Value", main="")
lines(x.value, T.test)
```

R-Code for the Change-Point test (10000 simulated samples)

```
#Reading data and fitting GARCH(1,1) model
set.seed(24)

data <- scan("data.txt")

dwj =0
dwj = data[length(data):1]
dwj.log <- log(dwj)
dwj.log.diff <- diff(dwj.log)

garch11 = garchFit(~garch(1,1), dwj.log.diff)
coef <- garch11@fit$coef[2:4]

x = 0
sigma = 0

sigma[1] = sqrt(coef[1]/(1-coef[2]-coef[3]))

for (k in 1:10000){

x = garchSim(model=list(omega=coef[1], alpha=coef[2], beta=coef[3]), n=2516)
xnew = x[1:2515]

g = garchFit(~garch(1,1), xnew)
g.coef = g@fit$coef[2:4]

n = length(dwj.log.diff)-1
  m1<-matrix(rep(c(g.coef[3]^(0:n),0),n+1), nrow=n+1, ncol=n+1)
m1[upper.tri(m1)]=0
```

```
mA=m1%*%(x[1:length(dwj.log.diff)]^2)

m21<-matrix(rep(c(0, g.coef[3]^(0:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m22<-matrix(rep(c(0, (1:n),0),n+1), nrow=n+1, ncol=n+1)
m2<-m21*m22

mB=m2%*%(x[1:length(dwj.log.diff)]^2)

m31<-matrix(rep(c(0,0, g.coef[3]^(0:(n-2)),0),n+1), nrow=n+1, ncol=n+1)
m32<-matrix(rep(c(0,0,(1:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m33<-matrix(rep(c(0,0,(2:n),0),n+1), nrow=n+1, ncol=n+1)
m3<-m31*m32*m33

mC=m3%*%(x[1:length(dwj.log.diff)]^2)

d11.vec = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma[1]^2), x, (length(x)-1), mA, mB, mC)$d1.vec
d.matrix = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma[1]^2), x, (length(x)-1), mA, mB, mC)$d2.matrix

d.vec=0
for(i in 1:length(d11.vec[,1])){
d.vec[i]=list(d11.vec[i,])
}

T.test = sapply(1:length(d.vec), Test, d.vec, d.matrix)

T[k] = mean(T.test)
}

T
```

R-Code to investigate the power of the test, based on the full data set.

```
set.seed(24)

data <- scan("data.txt")

dwj =0

dwj = data[length(data):1]
dwj.log <- log(dwj)
dwj.log.diff <- diff(dwj.log)

data1 = dwj.log.diff[1:1480]
```

```
data2 = dwj.log.diff[1481:length(dwj.log.diff)]

garch1 = garchFit(~garch(1,1),data1)
garch2 = garchFit(~garch(1,1),data2)

coef1 = garch1@fit$coef[2:4]
coef2 = garch2@fit$coef[2:4]

x1 = 0
sigma1 = 0

x2 = 0

sigma1[1] = sqrt(coef1[1]/(1-coef1[2]-coef1[3]))

for (k in 1:10000){
x1 = garchSim(model=list(omega=coef1[1], alpha=coef1[2], beta=coef1[3]), n=(length(data1)+1))
x2 = garchSim(model=list(omega=coef2[1], alpha=coef2[2], beta=coef2[3]), n=length(data2))

x = c(x1, x2)
xnew = c(x1[2:(length(data1)+1)], x2)

g = garchFit(~garch(1,1),xnew)
g.coef = g@fit$coef[2:4]

n = length(xnew)-1
  m1<-matrix(rep(c(g.coef[3]^(0:n),0),n+1), nrow=n+1, ncol=n+1)
m1[upper.tri(m1)]=0

mA=m1%*%(x[1:length(xnew)]^2)

m21<-matrix(rep(c(0, g.coef[3]^(0:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m22<-matrix(rep(c(0, (1:n),0),n+1), nrow=n+1, ncol=n+1)
m2<-m21*m22

mB=m2%*%(x[1:length(xnew)]^2)

m31<-matrix(rep(c(0,0, g.coef[3]^(0:(n-2)),0),n+1), nrow=n+1, ncol=n+1)
m32<-matrix(rep(c(0,0,(1:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m33<-matrix(rep(c(0,0,(2:n),0),n+1), nrow=n+1, ncol=n+1)
m3<-m31*m32*m33

mC=m3%*%(x[1:length(xnew)]^2)

d11.vec = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma1[1]^2), x, (length(x)-1), mA, mB, mC)$d1.vec
d.matrix = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma1[1]^2), x, (length(x)-1), mA, mB, mC)$d2.matrix
```

```
d.vec=0
for(i in 1:length(d11.vec[,1])){
d.vec[i]=list(d11.vec[i,])
}

T.test = sapply(1:length(d.vec), Test, d.vec, d.matrix)

T[k] = mean(T.test)
}
T
```

R-Code for bootstrapping test results (10000 bootstrapping samples)

```
#Reading data and fitting GARCH(1,1) model
set.seed(24)

data <- scan("data.txt")

dwj = 0
dwj = data[length(data):1]
dwj.log <- log(dwj)
dwj.log.diff <- diff(dwj.log)

garch11 = garchFit(~garch(1,1),dwj.log.diff)
coef <- garch11@fit$coef[2:4]

x = 0
sigma = 0
error =0

sigma[1] = sqrt(coef[1]/(1-coef[2]-coef[3]))

x0 = sqrt(coef[1]/(1-coef[2]-coef[3]))

x= c(x0, dwj.log.diff)

for (i in 2:(length(dwj.log.diff)+1)){
sigma[i] <- sqrt(coef[1]+(coef[2]*x[i-1]^2)+(coef[3]*sigma[i-1]^2))
error[i-1] = x[i]/sigma[i]
}

sigma.boot=0
sigma.boot[1] = sqrt(coef[1]/(1-coef[2]-coef[3]))
```

```
for (k in 1:1){
x.boot = 0
x.boot[1] = sqrt(coef[1]/(1-coef[2]-coef[3]))
for (i in 2:(length(dwj.log.diff)+1)){
sigma.boot[i] <- sqrt(coef[1]+(coef[2]*x.boot[i-1]^2)+(coef[3]*sigma.boot[i-1]^2))
x.boot[i] <- rnorm(1, mean(error), sqrt(var(error)))*sigma.boot[i]
}

g = garchFit(~garch(1,1),x.boot)
g.coef = g@fit$coef[2:4]

n = length(dwj.log.diff)-1
  m1<-matrix(rep(c(g.coef[3]^(0:n),0),n+1), nrow=n+1, ncol=n+1)
m1[upper.tri(m1)]=0

mA=m1%*%(x.boot[1:length(dwj.log.diff)]^2)

m21<-matrix(rep(c(0, g.coef[3]^(0:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m22<-matrix(rep(c(0, (1:n),0),n+1), nrow=n+1, ncol=n+1)
m2<-m21*m22

mB=m2%*%(x.boot[1:length(dwj.log.diff)]^2)

m31<-matrix(rep(c(0,0, g.coef[3]^(0:(n-2)),0),n+1), nrow=n+1, ncol=n+1)
m32<-matrix(rep(c(0,0,(1:(n-1)),0),n+1), nrow=n+1, ncol=n+1)
m33<-matrix(rep(c(0,0,(2:n),0),n+1), nrow=n+1, ncol=n+1)
m3<-m31*m32*m33

mC=m3%*%(x.boot[1:length(dwj.log.diff)]^2)

d11.vec = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma.boot[1]^2), x.boot, (length(x.boot)-1),
 mA, mB, mC)$d1.vec
d.matrix = Deriv(g.coef[1], g.coef[3], g.coef[2], (sigma.boot[1]^2), x.boot, (length(x.boot)-1),
 mA, mB, mC)$d2.matrix

d.vec=0
for(i in 1:length(d11.vec[,1])){
d.vec[i]=list(d11.vec[i,])
}

T.test = sapply(1:length(d.vec), Test, d.vec, d.matrix)

T[k] = max(T.test)
}
T
```

R-Code for Change-Point test for the first split.

```
data <- scan("data.txt")

dwj = 0
dwj = data[length(data):1]
dwj.log <- log(dwj)
dwj.log.diff <- diff(dwj.log)

data1 = dwj.log.diff[1:1480]
data2 = dwj.log.diff[1481:length(dwj.log.diff)]

garch1 = garchFit(~garch(1,1), data1)
garch2 = garchFit(~garch(1,1), data2)

coef1 = garch1@fit$coef[2:4]
coef2 = garch2@fit$coef[2:4]

x10 = sqrt(coef1[1]/(1-coef1[2]-coef1[3]))
sigma10_2 = (coef1[1]/(1-coef1[2]-coef1[3]))

newx1 = c(x10, data1)

x20 = sqrt(coef2[1]/(1-coef2[2]-coef2[3]))
sigma20_2 = (coef2[1]/(1-coef2[2]-coef2[3]))

newx2 = c(x20, data2)
n1 = length(data1)-1
m1_1<-matrix(rep(c(coef1[3]^(0:n1),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_1[upper.tri(m1_1)]=0

m1A=m1_1%*%(newx1[1:length(data1)]^2)

m1_21<-matrix(rep(c(0, coef1[3]^(0:(n1-1)),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_22<-matrix(rep(c(0, (1:n1),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_2<-m1_21*m1_22

m1B=m1_2%*%(newx1[1:length(data1)]^2)

m1_31<-matrix(rep(c(0,0, coef1[3]^(0:(n1-2)),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_32<-matrix(rep(c(0,0,(1:(n1-1)),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_33<-matrix(rep(c(0,0,(2:n1),0),n1+1), nrow=n1+1, ncol=n1+1)
m1_3<-m1_31*m1_32*m1_33

m1C=m1_3%*%(newx1[1:length(data1)]^2)

n2 = length(data2)-1
```

```
m2_1<-matrix(rep(c(coef2[3]^(0:n2),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_1[upper.tri(m2_1)]=0

m2A=m2_1%*%(newx2[1:length(data2)]^2)

m2_21<-matrix(rep(c(0, coef2[3]^(0:(n2-1)),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_22<-matrix(rep(c(0, (1:n2),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_2<-m2_21*m2_22

m2B=m2_2%*%(newx2[1:length(data2)]^2)

m2_31<-matrix(rep(c(0,0, coef2[3]^(0:(n2-2)),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_32<-matrix(rep(c(0,0,(1:(n2-1)),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_33<-matrix(rep(c(0,0,(2:n2),0),n2+1), nrow=n2+1, ncol=n2+1)
m2_3<-m2_31*m2_32*m2_33

m2C=m2_3%*%(newx2[1:length(data2)]^2)

d11.vec = Deriv(coef1[1], coef1[3], coef1[2], sigma10_2, newx1, length(data1), m1A, m1B, m1C)$d1.vec
d1.matrix = Deriv(coef1[1], coef1[3], coef1[2], sigma10_2, newx1, length(data1), m1A, m1B, m1C)$d2.matrix

d.vec1=0
for(i in 1:length(d11.vec[,1])){
d.vec1[i]=list(d11.vec[i,])
}

d21.vec = Deriv(coef2[1], coef2[3], coef2[2], sigma20_2, newx2, length(data2), m2A, m2B, m2C)$d1.vec
d2.matrix = Deriv(coef2[1], coef2[3], coef2[2], sigma20_2, newx2, length(data2), m2A, m2B, m2C)$d2.matrix

d.vec2=0
for(i in 1:length(d21.vec[,1])){
d.vec2[i]=list(d21.vec[i,])
}

T.test1 = sapply(1:length(d.vec1), Test, d.vec1, d1.matrix)
T.test1

max(T.test1)
mean(T.test1)

T.test2 = sapply(1:length(d.vec2), Test, d.vec2, d2.matrix)
T.test2

max(T.test2)
mean(T.test2)
```

```
x_axis = function(index=1, vec){
return(index/(length(vec)))
}

xvec1=c(rep(1, times=length(d.vec1)))
x.value1 = sapply(1:length(xvec1), x_axis, xvec1)

plot(x.value1, T.test1, xlab="t", ylab="Test Statistic Value", main=" ")
lines(x.value1, T.test1)

xvec2=c(rep(1, times=length(d.vec2)))
x.value2 = sapply(1:length(xvec2), x_axis, xvec2)

plot(x.value2, T.test2, xlab="t", ylab="Test Statistic Value", main="")
lines(x.value2, T.test2)
```

Bibliography

[1] Aue, A., Berkes, I., and Horváth, L. (2006). Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* **12**, 583–608.

[2] Berkes, I., Horváth, L., and Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p, q) models. *Statistics & Probability Letters* **70**, 263–273.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[4] Coulson, N.E., and Robins, R.P. (1985). Aggregate economic activity and the variance of inflation: Another look. *Economics Letters* **17**, 71-75.

[5] Domowitz, I., and Hakkio, C.S. (1985). Conditional variance and the risk premium in the foreign exchange market. *Journal of International Economics* **19**, 47-66.

[6] Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987-1008.

[7] Engle, R.F. (1983). Estimates of the variance of U.S. inflation based on the ARCH model. *Journal of Money Credit and Banking* **15**, 286-301.

[8] Engle, R.F. and Kraft, D. (1983). Multiperiod forecast erroe varainces of inflation estimated from ARCH models, in: A.Zellner, ed., Applied time series analysis of ecnomic data (Bureau of the Census, Washington, DC), 293-302.

[9] Engle, R.F, Lilien, D., and Robins, R. (1985). Estimation of time varying risk premium in the term structure, Discussion paper 85-117 (University of California, San Diego, CA).

[10] Furstenberg, H., and Kesten, H. (1960). Products of random matrices. *The Annals of Mathematical Statistics* **31**, 457-469.

[11] Kiefer, J. (1959). K-Sample analogues of the Kolmogorov-Smirnov and Cramér-v.Mises tests. *The Annals of Mathematical Statistics* **30**, 420-447.

[12] Mikosch, T. and Stărică, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *The Annals of Statistics* **28**, 1427–1451.

[13] Shiryaev, A.N. (1989). *Graduate Texts in Mathematics*, second edition. Springer-Verlag, New York, 254.

[14] Weiss, A.A. (1984). ARMA models with ARCH errors. *Journal of Time Series Analysis* **5**, 129-143.